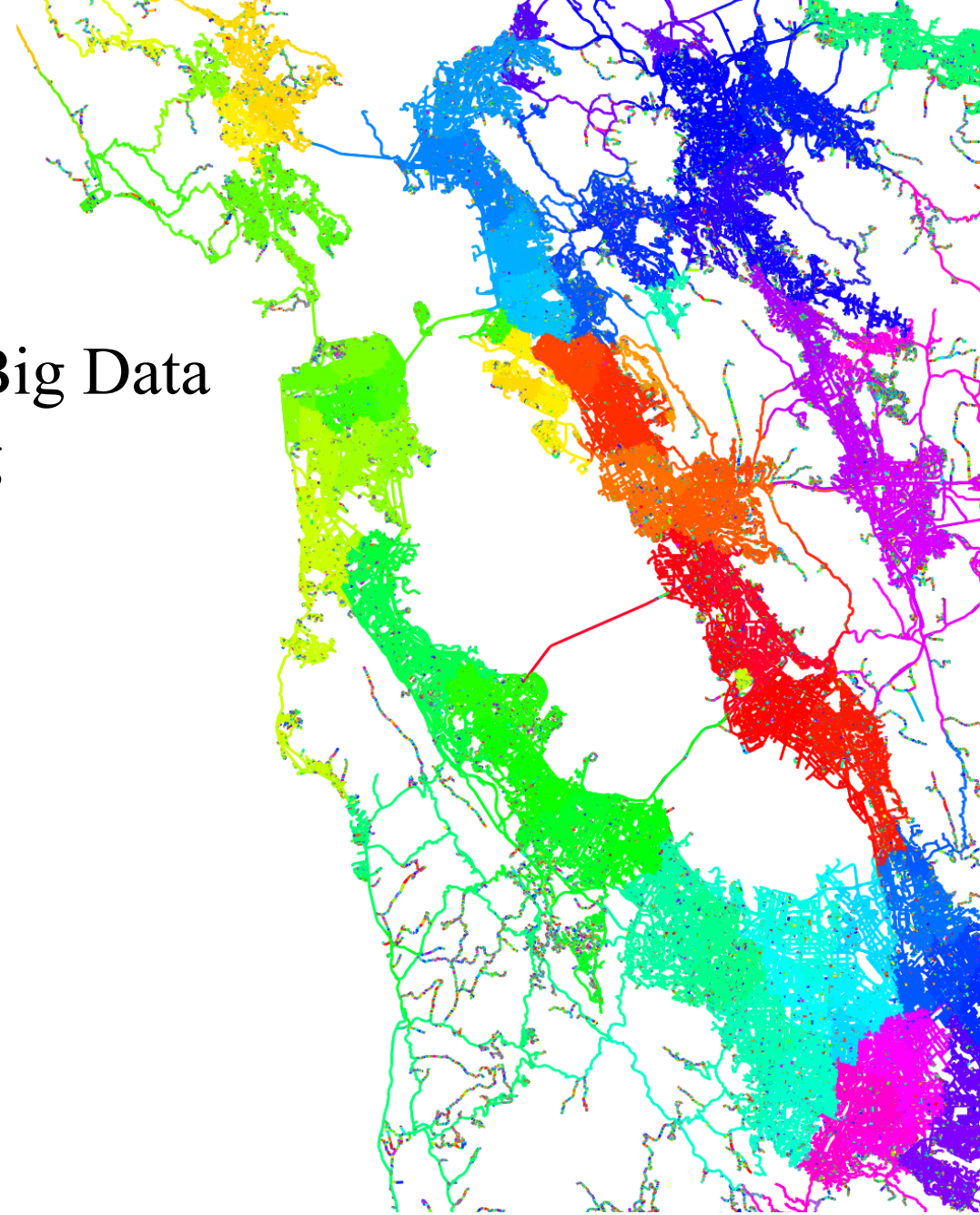


# 2020 DOE Vehicle Technologies Office Annual Merit Review

## High-Performance Computing (HPC) and Big Data Solutions for Mobility Design and Planning

Jane Macfarlane  
Lawrence Berkeley National Laboratory  
June 12, 2020

Project ID: eems037



# Overview

## TIMELINE

- Start: October 2017
- End: September 2020
- 80 % complete

## BUDGET

- Total project funding
- \$3.9 M / 3 years
- \$1.3 M per year / 2 Labs

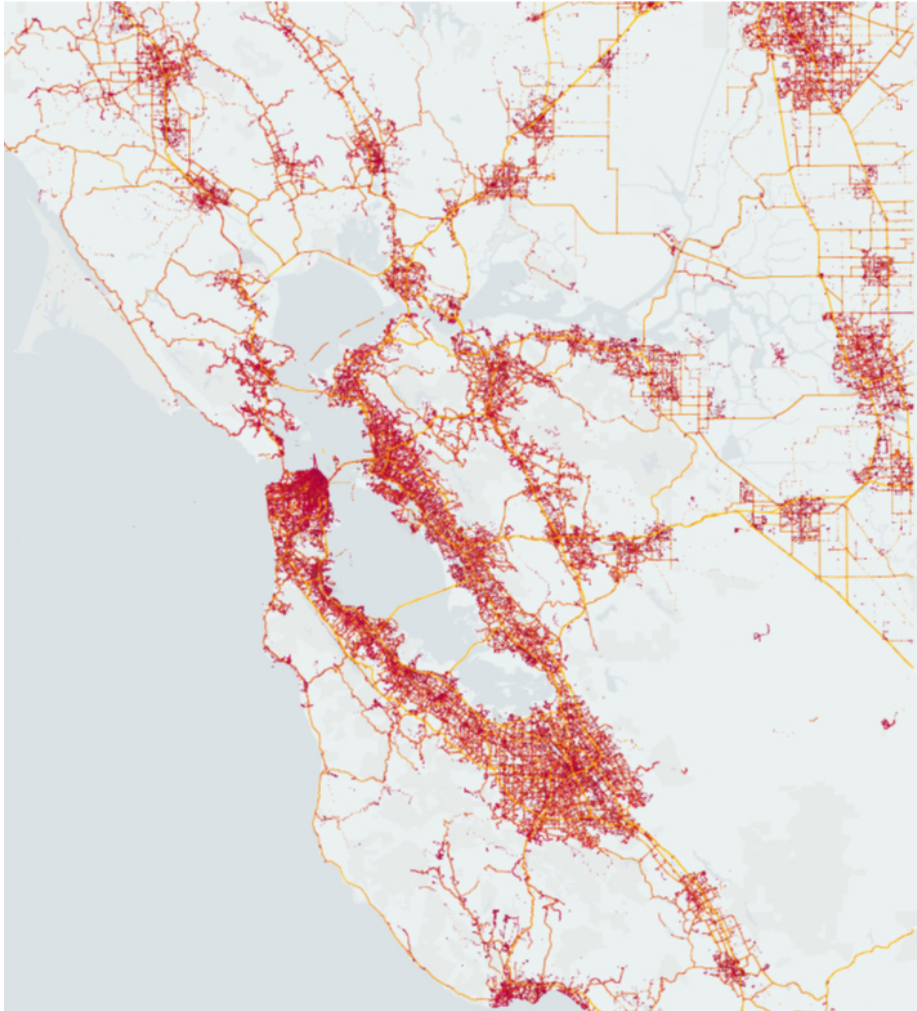
## PARTNERS

- CalTrans Connected Corridors
- HERE Technologies

## BARRIERS

- Metropolitan scale networks are too complex to model in reasonable compute time.
- Sensors for capturing dynamics provide limited view and are difficult to mine for relevant information.
- Optimization of energy, travel time and mobility across complex networks has yet to be accomplished for real-world metropolitan scale networks.

# Relevance and Project Objectives



## Overall Objective:

- Develop HPC tools to **rapidly model large-scale transportation networks** using real-world, near real-time data. Integrate energy, travel time and mobility measures to determine optimization opportunities.

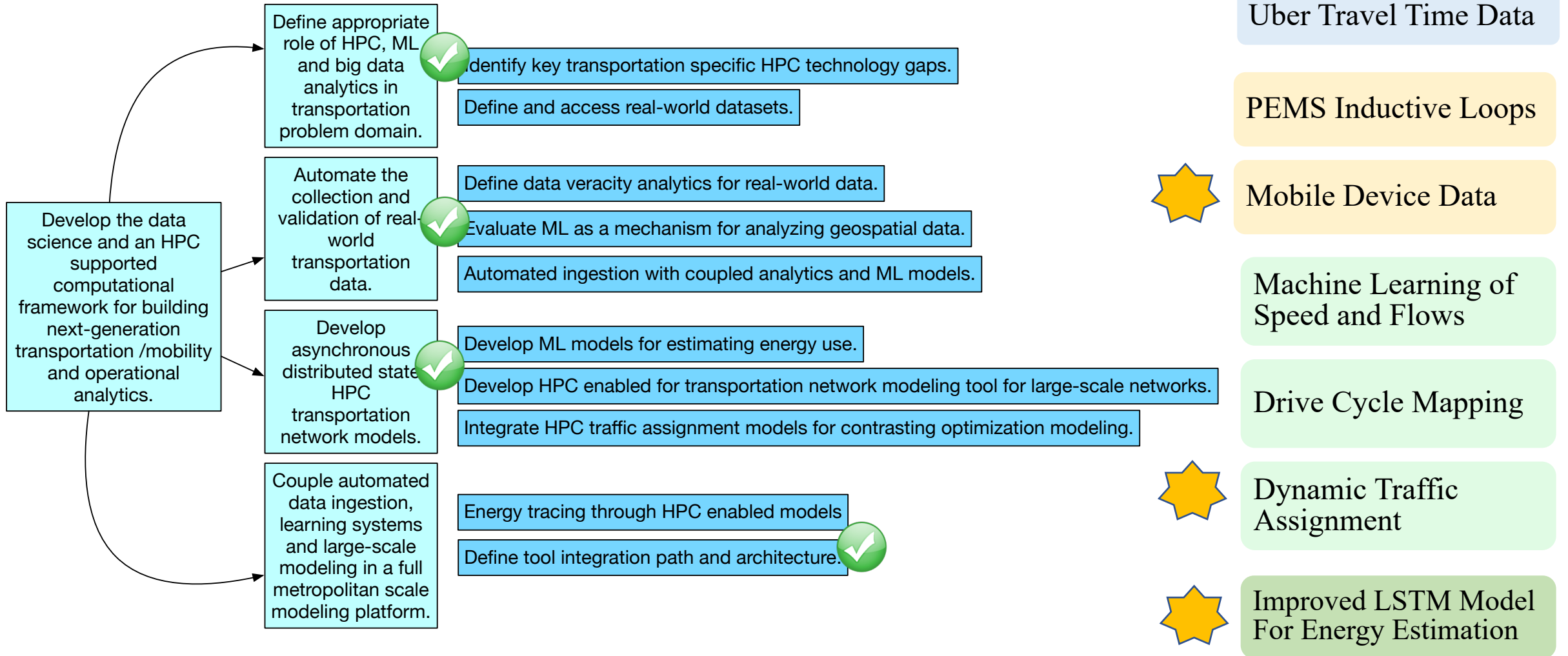
## Objectives this Period:

- Extend **validation** methods to include real-world data
- **Cross validate estimates of the energy cost** and productivity loss of congestion **using data-driven approach**.
- Integrate machine learning of real-world sensors into link models.


## Impact:

- Develop new **active control ideas for connected vehicles** that will optimize energy, travel time and mobility for normal traffic conditions and networks under stress.

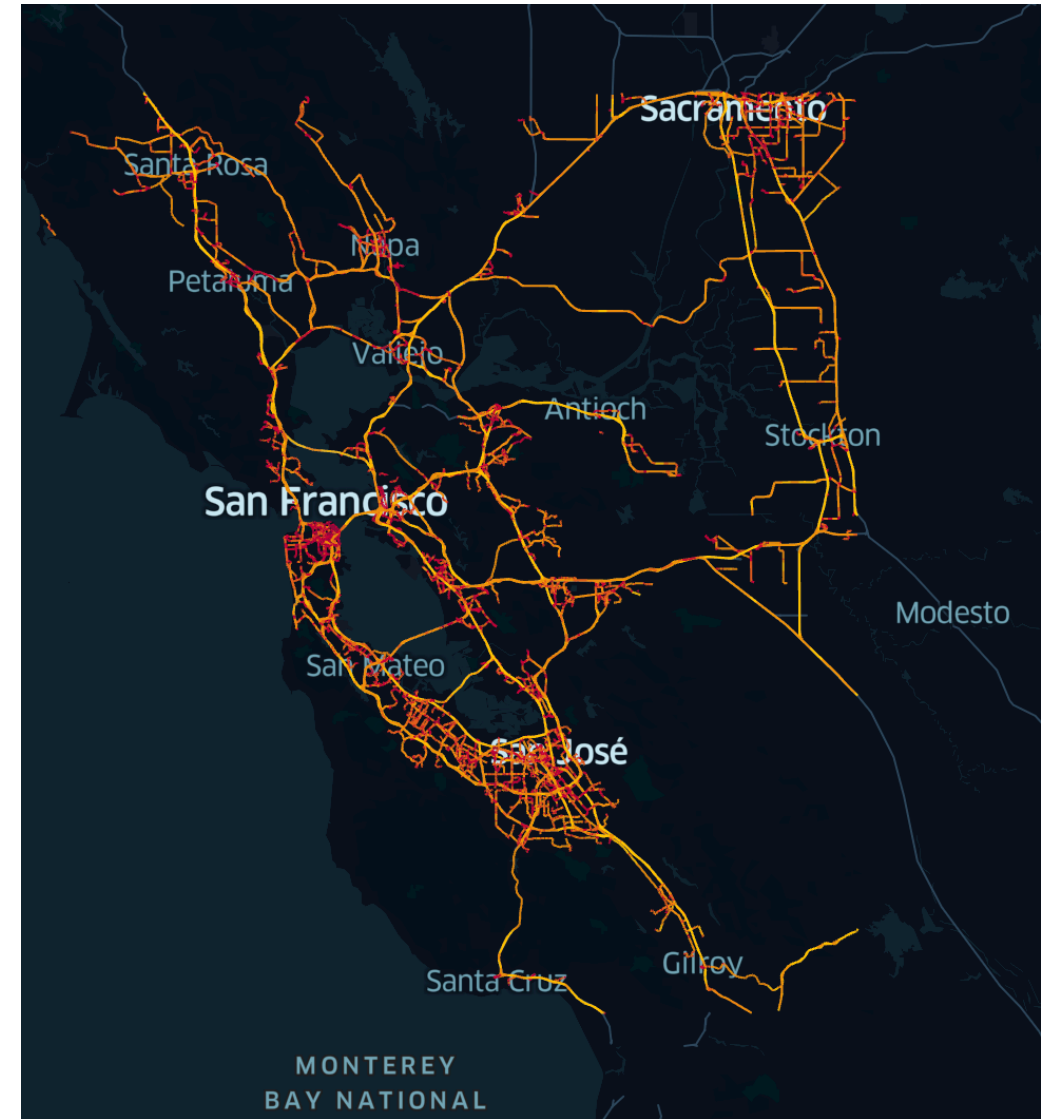
# Project Goals



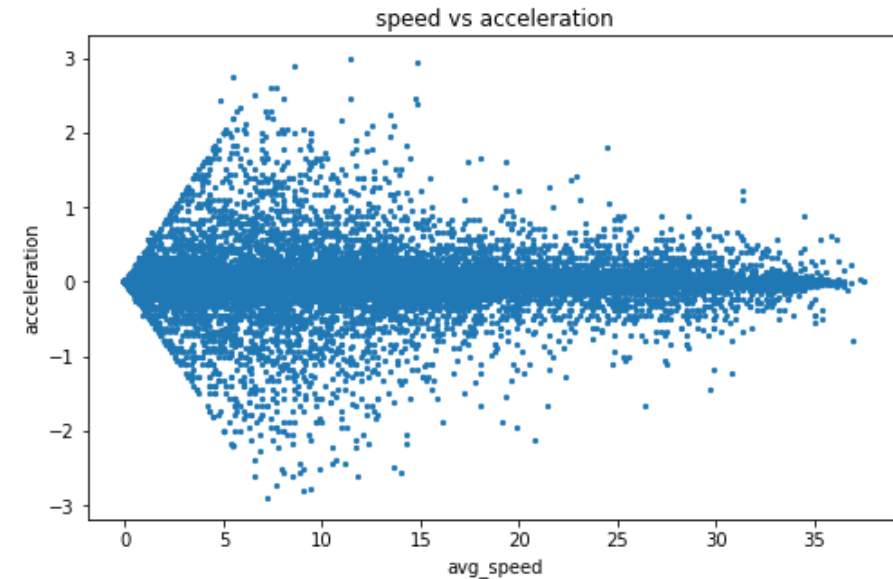
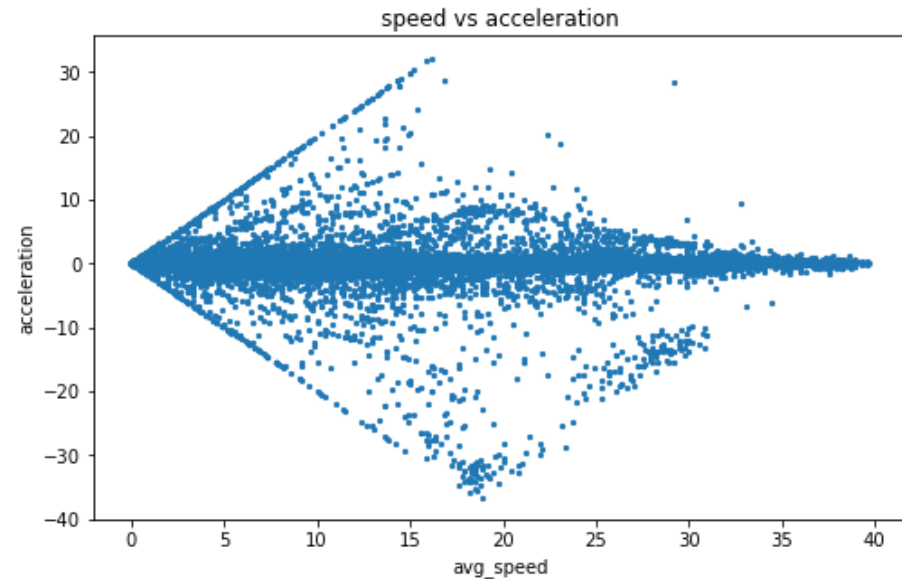
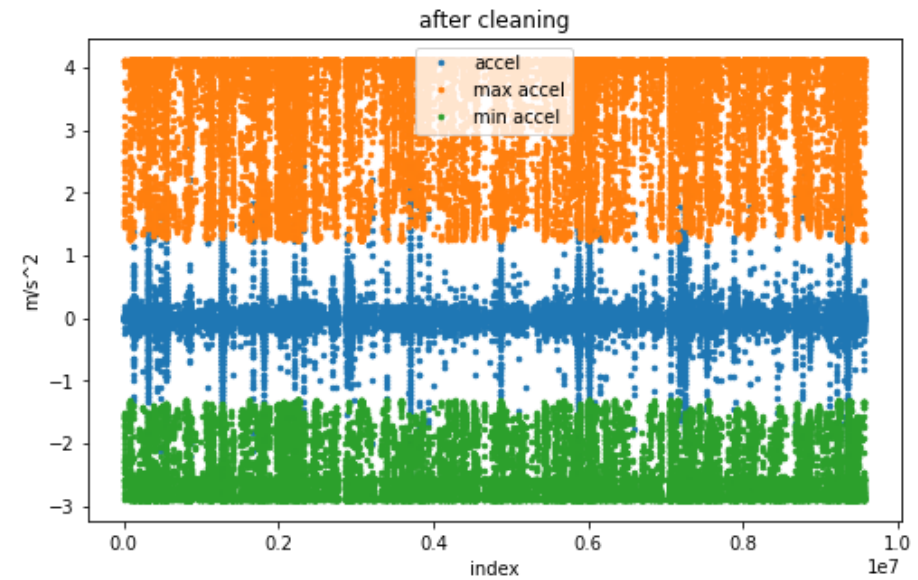
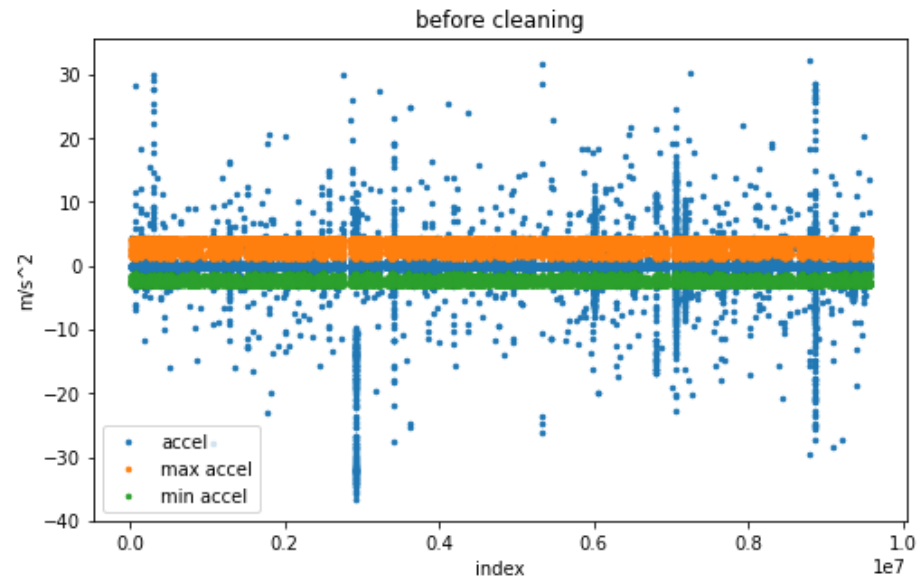
# Go/NoGo Milestones

Define appropriate role of HPC, ML and big data analytics in transportation problem domain.	Defined goals for developing metropolitan scale modeling. Alliances with SF and San Jose.	Identified traffic assignment optimization research for integration. <del>Collaboration established with Dallas Ft Worth/TTI.</del>		Continuing
Automate the collection and validation of real-world transportation data.		Go/NoGo - Demonstrated good modeling of speed and flow with DCRNN with automated ingestion of loop detectors.	Use of probe data as virtual sensors to augment current loop detectors geospatial range.	On Track
		Developed data driven ML models for estimating energy consumption.	Integrated energy estimation. 	On Track
Develop large-scale HPC enabled transportation network models.	Go/NoGo - Mobiliti model developed that models metropolitan scale network with compute time < 1 minute.	Go/NoGo - Mobiliti model developed that models metropolitan scale network capability of dynamic routing.	Investigate Active Control methods focused on reduction of energy and increased mobility.	On Track
Couple automated data ingestion, learning systems and large-scale modeling in a full metropolitan scale modeling platform.			Go/NoGo - Integration of ML models into the link dynamic models in Mobiliti.	

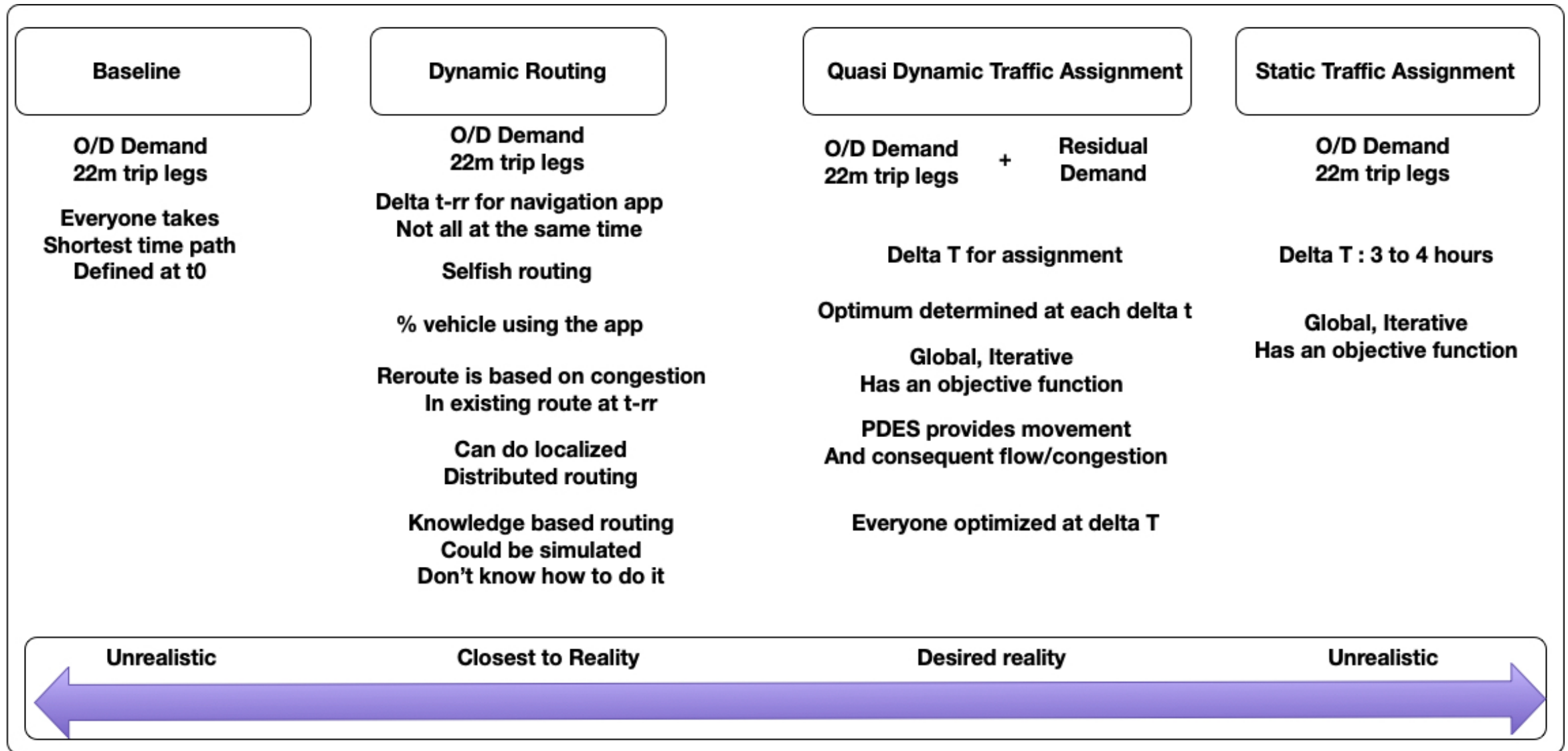
# Raw Data Cleaning : GPS Veracity



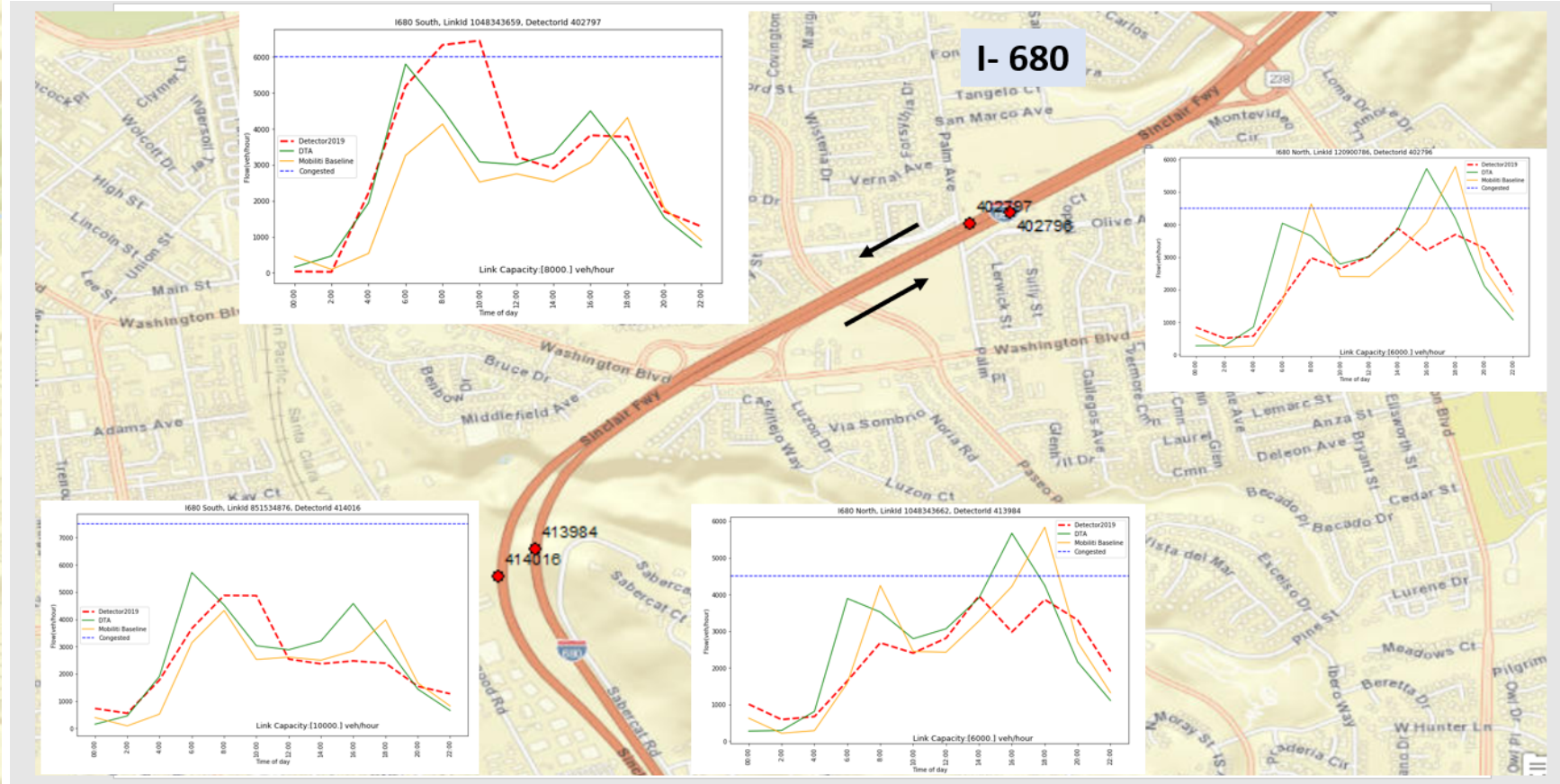
# Raw Data Cleaning : Physical Properties



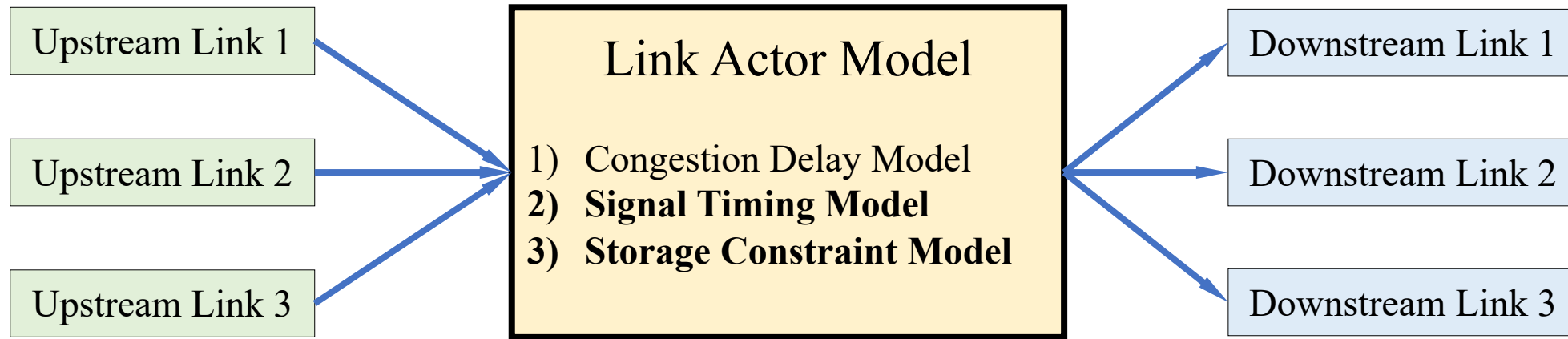
# Validation: Context



# Validation : Detector Data

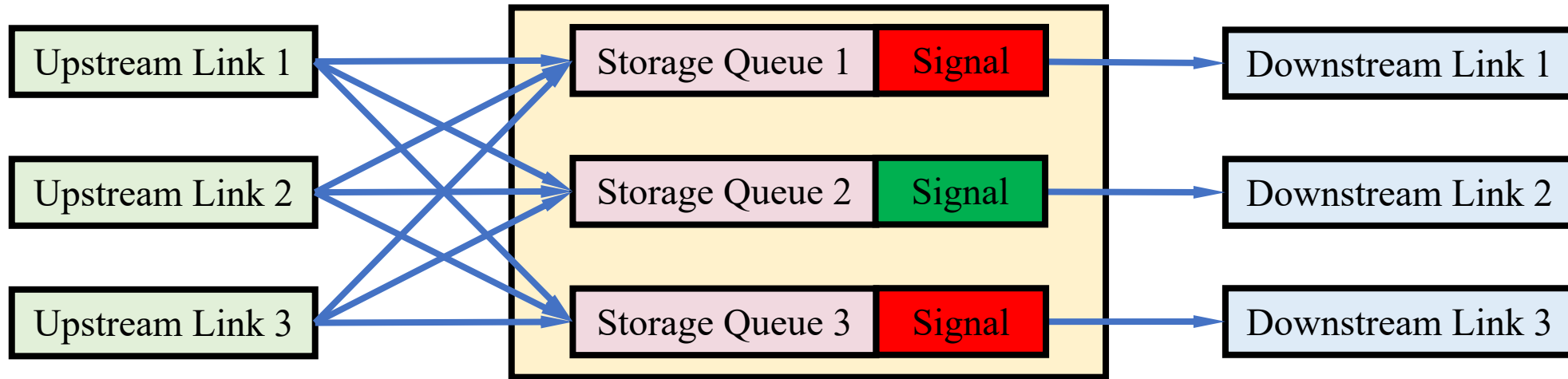


# Link Model Enhancements



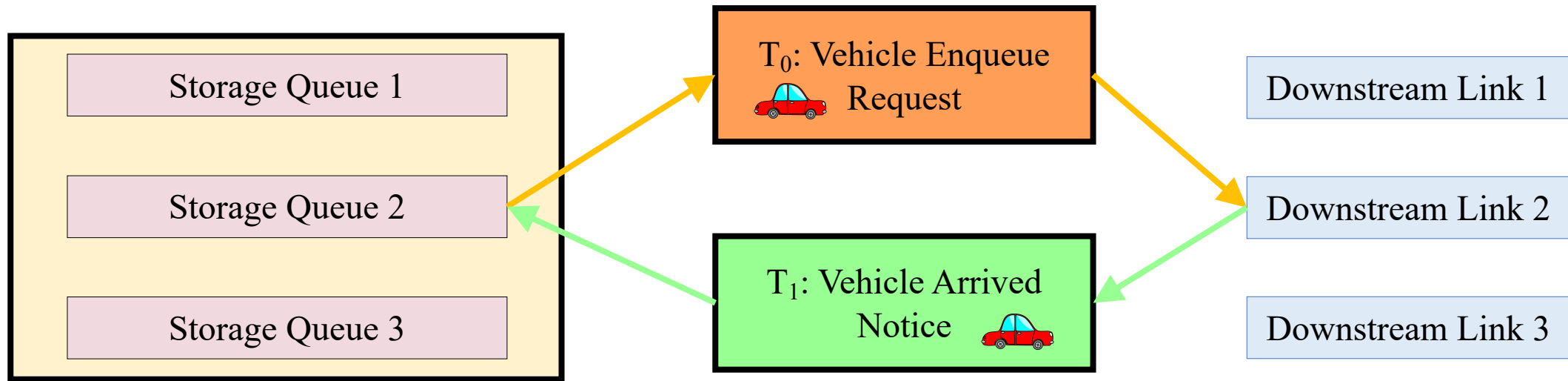
- Simulated link actor has several behavior models to capture the impact of various system dynamics:
  - Congestion delay
  - **Signal timing effects**
  - **Storage capacity constraints**

# Signal Timing Model



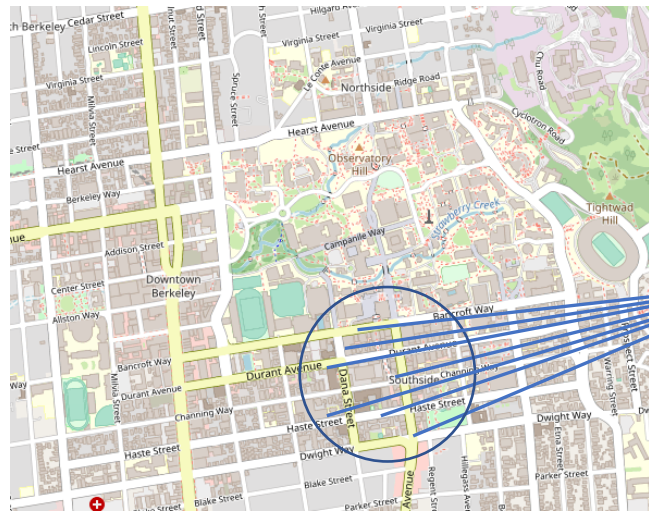
- Separate storage queue for each downstream link inside each actor
- Associate signal timing model with each outgoing link
  - Fully parameterized with offset and duration of the green signal phase for each downstream link
  - Parameters may be changed dynamically to experiment with adaptive signal timing control
  - Each green signal phase has a discrete number of time slots that vehicles may be assigned to transition to downstream link

# Storage Capacity Model



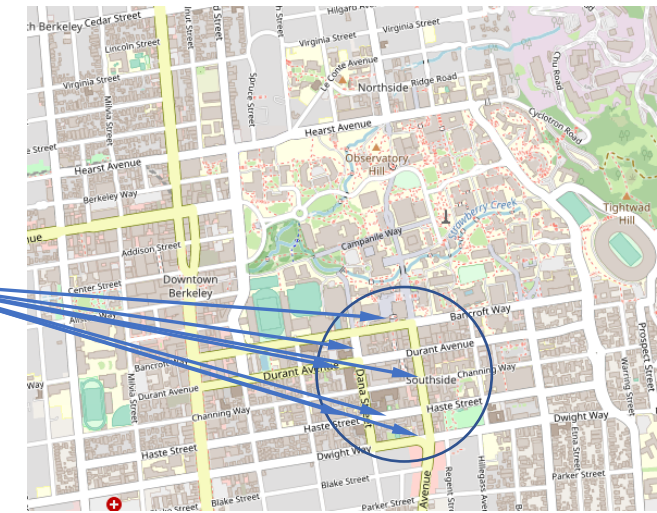
- Upstream link sends enqueue **request** with a minimum time  $T_0$  that vehicle may transition to downstream link
- Downstream link sends **notice** at time  $T_1$  that transition has actually occurred only when there is available storage capacity
- $T_0$  and  $T_1$  must both also obey signal timing constraints
- This mechanism captures spillback when congestion occurs downstream

# DCRNN Integration into Link Model (in-progress)



Aggregate link densities

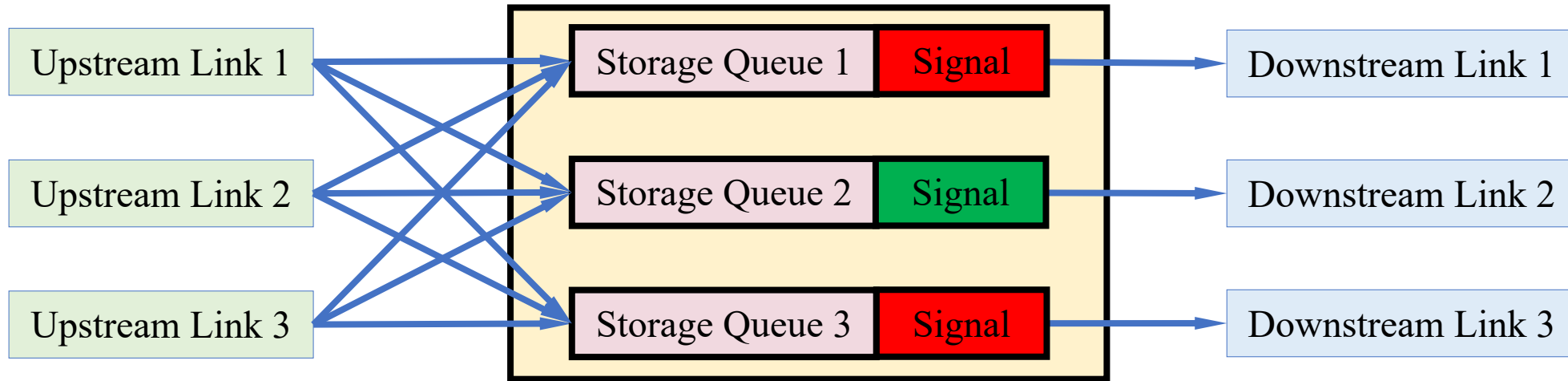
DCRNN Link  
Controller Agent



Report link speed estimates

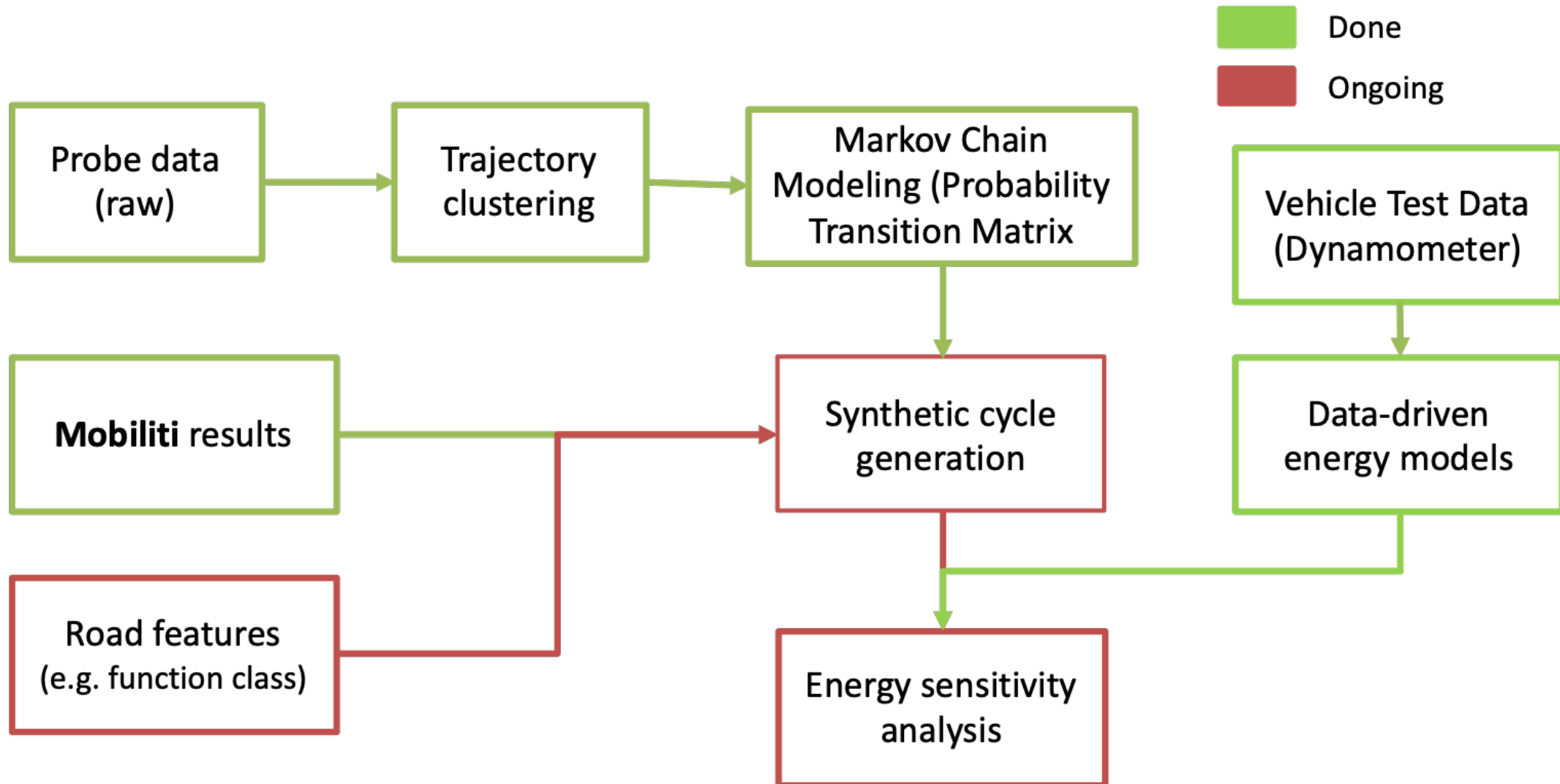
- DCRNN algorithm uses link density data from a neighborhood of links to predict link speeds
  - Data-driven approach to estimating link speeds and traversal times
- Will instantiate multiple DCRNN link controller agents in Mobiliti
  - Aggregate model inputs: input density data from a neighborhood of links
  - Periodically query pre-trained DCRNN model with aggregated data
  - Send link delay model parameter updates to link actors in the neighborhood

# Updated Link Model Summary

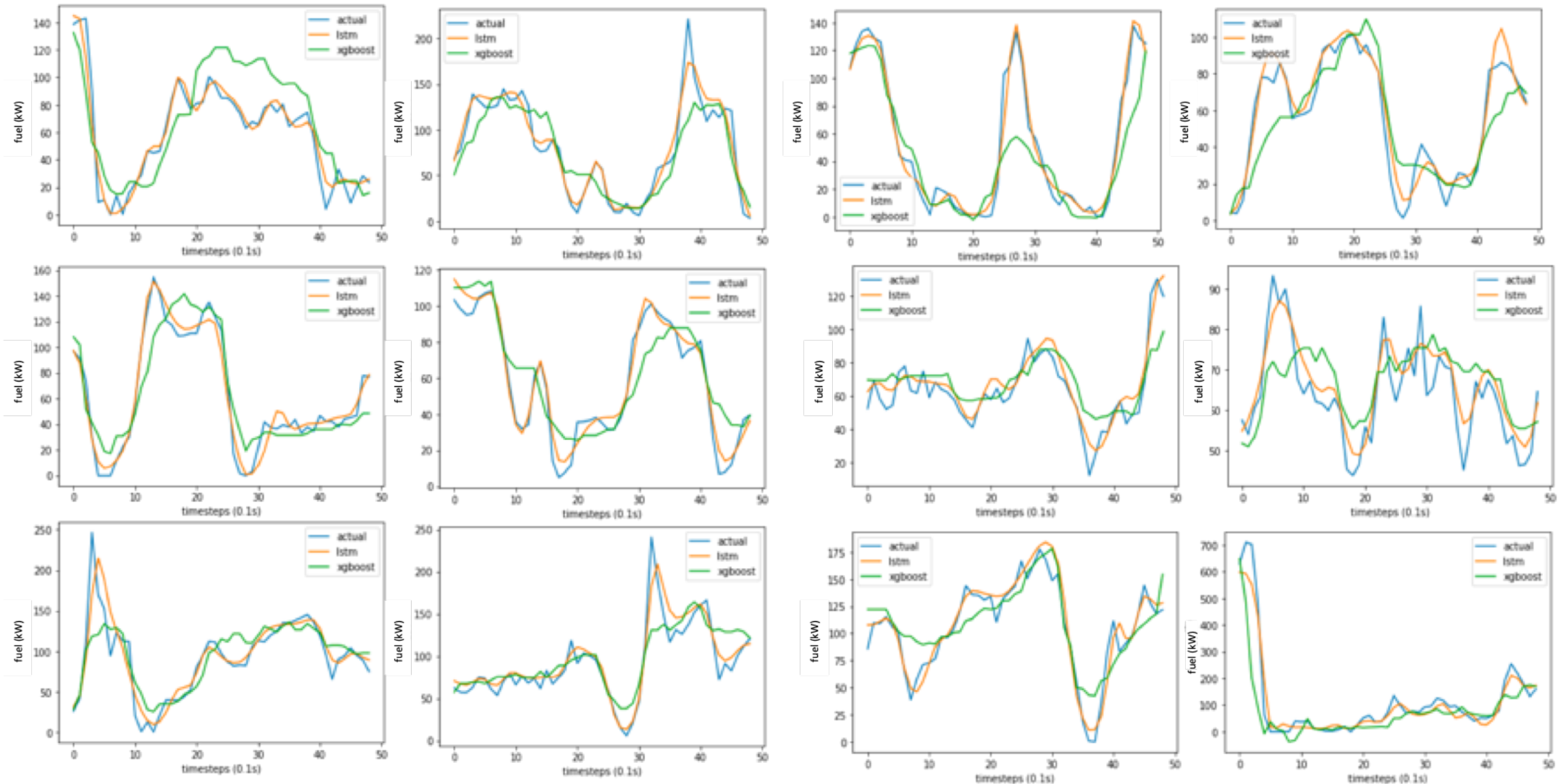


- Computation steps when each vehicle arrives at link actor:
  - Compute congested link traversal time with link delay model (e.g. DCRNN)
  - Compute additional delay due to waiting for next available green signal phase and assign vehicle to a time slot
  - Send vehicle enqueue request with preliminary transition time to downstream link
  - Downstream link will reply with actual transition time when it has free capacity

# Data Driven Energy Evaluation: WorkFlow



# Energy Prediction with a Data Driven Approach



# Response to Previous Comments

The reviewer noted that the proposed future work includes validation and integration. These are a natural extension of the current accomplishments.

- *Validation and integration of the real-world sensed data was a key focus of our work this year.*

The reviewer noted that this is very relevant as sitting in traffic and idling, otherwise known as congestion, is a major contributor to greenhouse gas emissions. If this project can anticipate and adjust traffic patterns to save energy and time with a byproduct of lower emissions, it is successful.

- *We thank the reviewer for supporting our mission to improve the quality of life in our cities through the reduction of energy used in congestion.*

Among the activities planned for future, it was not clear to the reviewer whether an effective active control algorithm can be thoroughly studied, even though it is going to be the major use of the developed platform.

- *While a difficult endeavor, we believe that with new artificial intelligence techniques and emerging technologies in edge computing in the cellular networks coupled with IoT in the infrastructure may provide significant opportunities for system level control. This simulation capability can offer a test bed for these new research ideas in which simple improvements could result in major congestion relief.*

# Collaboration and Coordination



UC Berkeley | ITS/PATH

Connected Corridors Program



# Challenges and Proposed Future Research

- External effects on link dynamics
- Surrogate models for codifying network constraints
- Traffic network characterization methods based on network connectivity, traffic dynamics, and spatiotemporal patterns
- Further develop and evolve data and model fusion methods for combining probe data with inductive loop data
- Develop mechanisms for understanding how to extract semantic knowledge from real-world sensor data to predict driver behavior
- Develop traffic-forecasting-specific uncertainty quantification approaches to assess uncertainty in the prediction

Any proposed future work is subject to change based on funding levels.

# Summary

- Developed data veracity pipeline for dealing with mobile device data
- Created large, cleaned data sets of trajectories for input to LSTM energy predictor
- Used external data from Uber and infrastructure embedded detectors for validation of simulation and Dynamic Traffic Assignment
- Added Storage and Signals to the Link Actor – maintained compute time
- Began integration of DCRNN link speed and flow predictor into Link Actor
- Improved and validated LSTM energy prediction

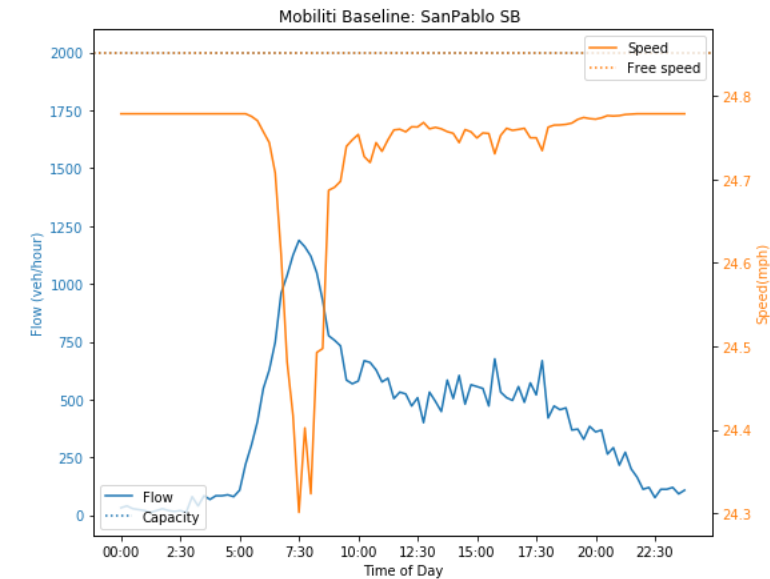
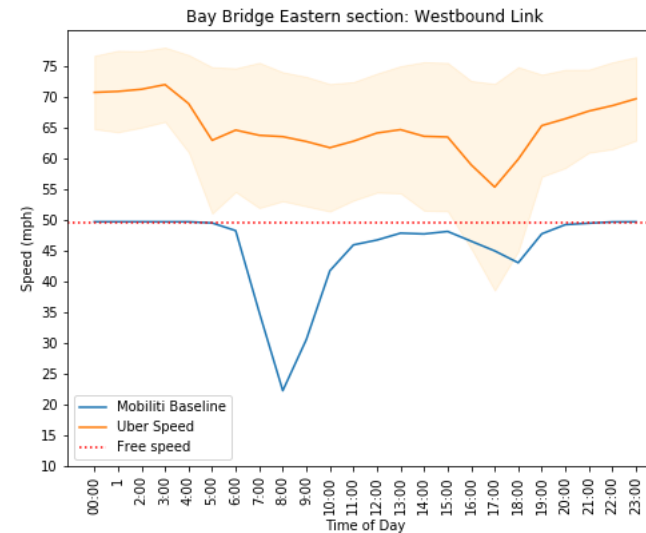
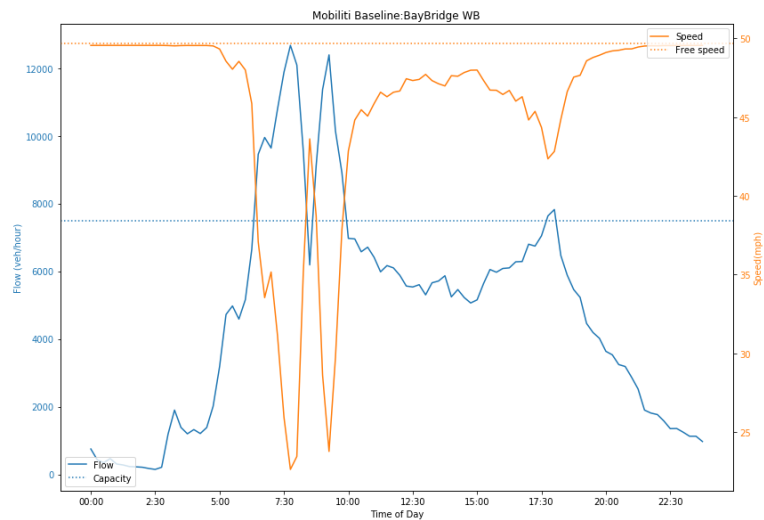
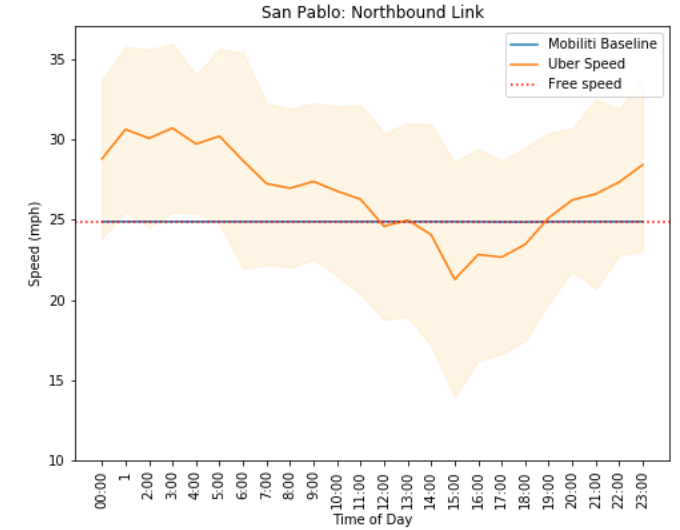
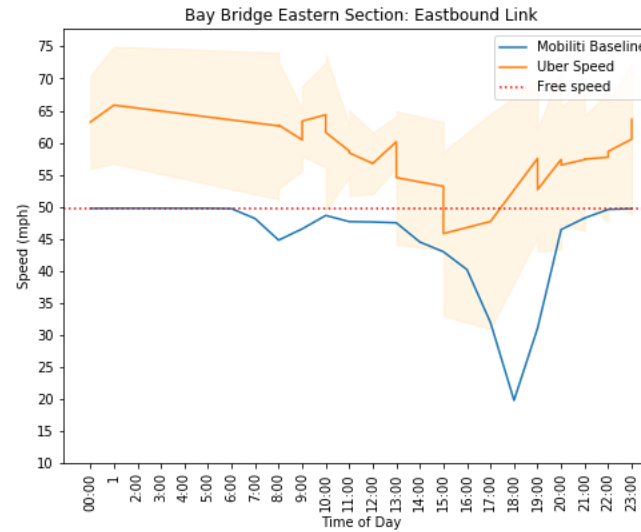
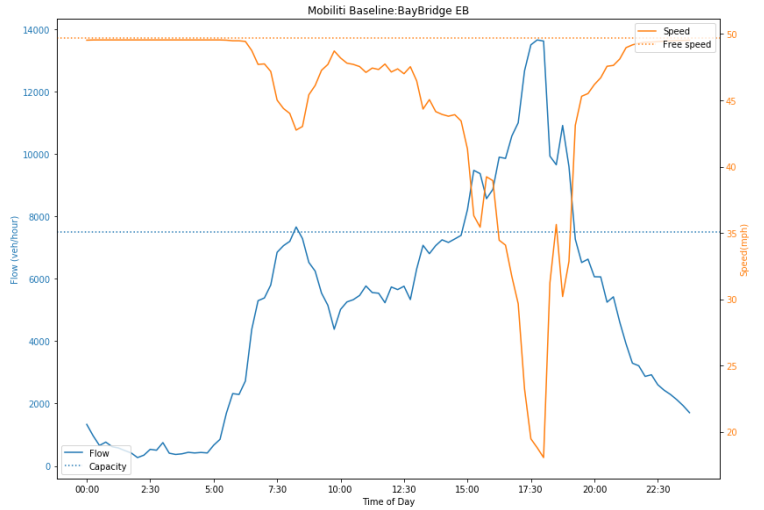
City	Road Miles	Area (sq miles)	Mobility VMT	Mobility Fuel (gal)	DTA VMT	DTA Fuel (gal)
San Francisco	1,951	46	6,323,569	242,051	6,268,216	238,651
San Jose	4,619	186	15,133,684	753,490	15,064,803	751,921
CC Corridor	9,861	440	53,066,175	2,416,471	52,192,435	2,410,732
Los Angeles	14,070	481	86,429,615	3,696,230	83,085,980	3,634,960

## Preliminary Results

Fleet Mix Assumption:  
40% compact vehicle  
40% mid size vehicle  
20% heavy duty truck

# Technical Backup

# Uber Validation : Bridge Speeds and Flows



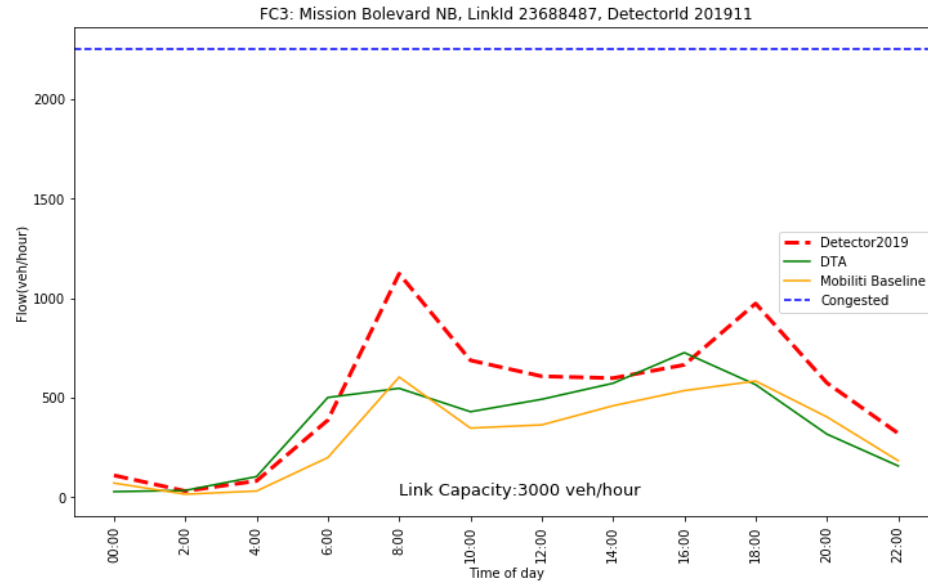
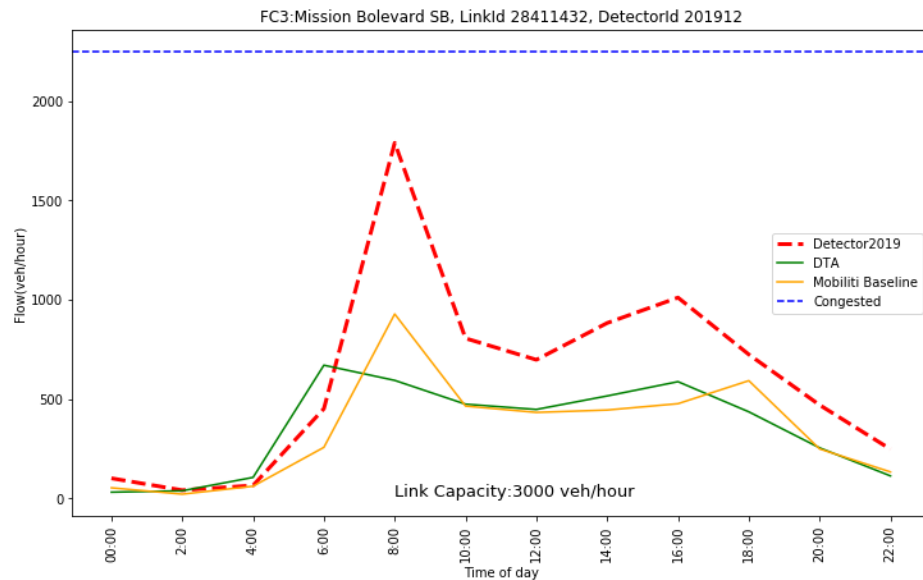
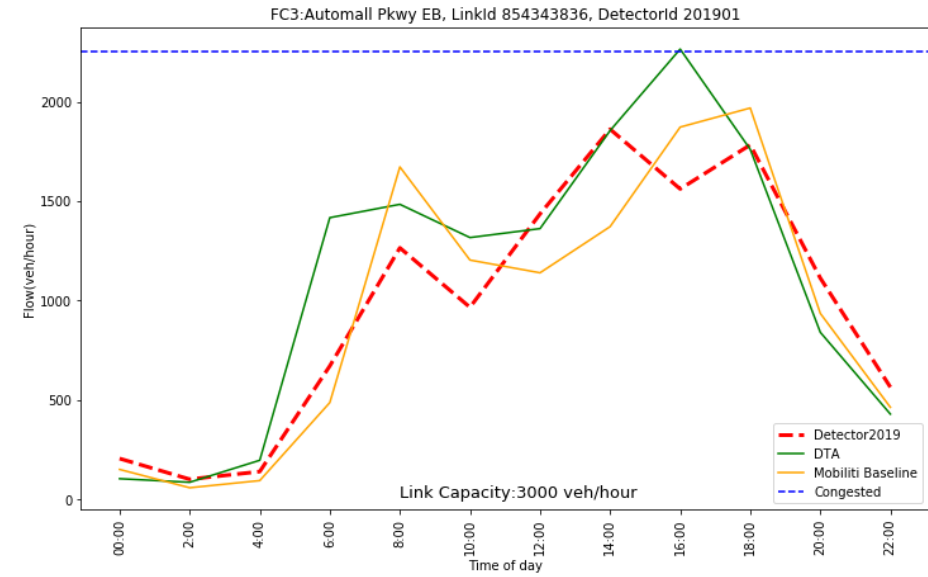
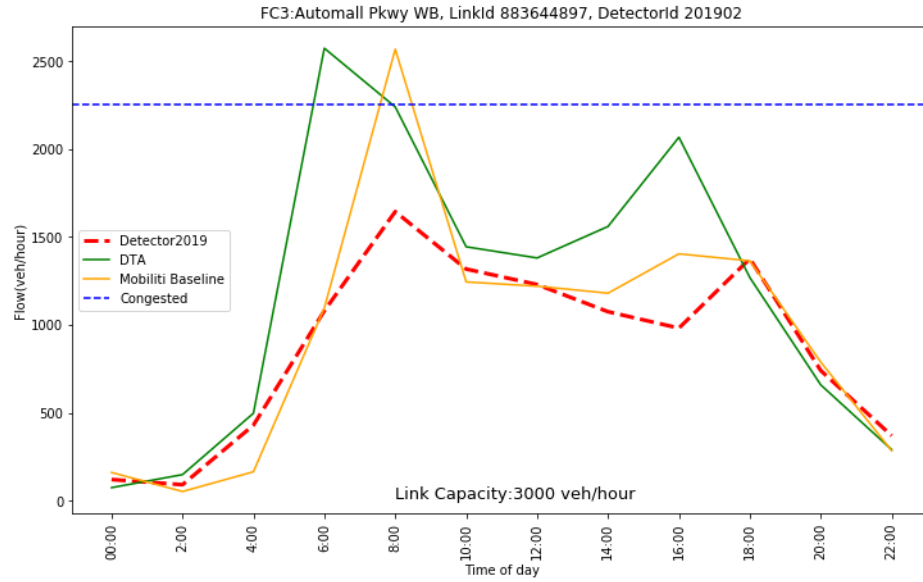
# Validation : Daily Bridge Counts

Mobility

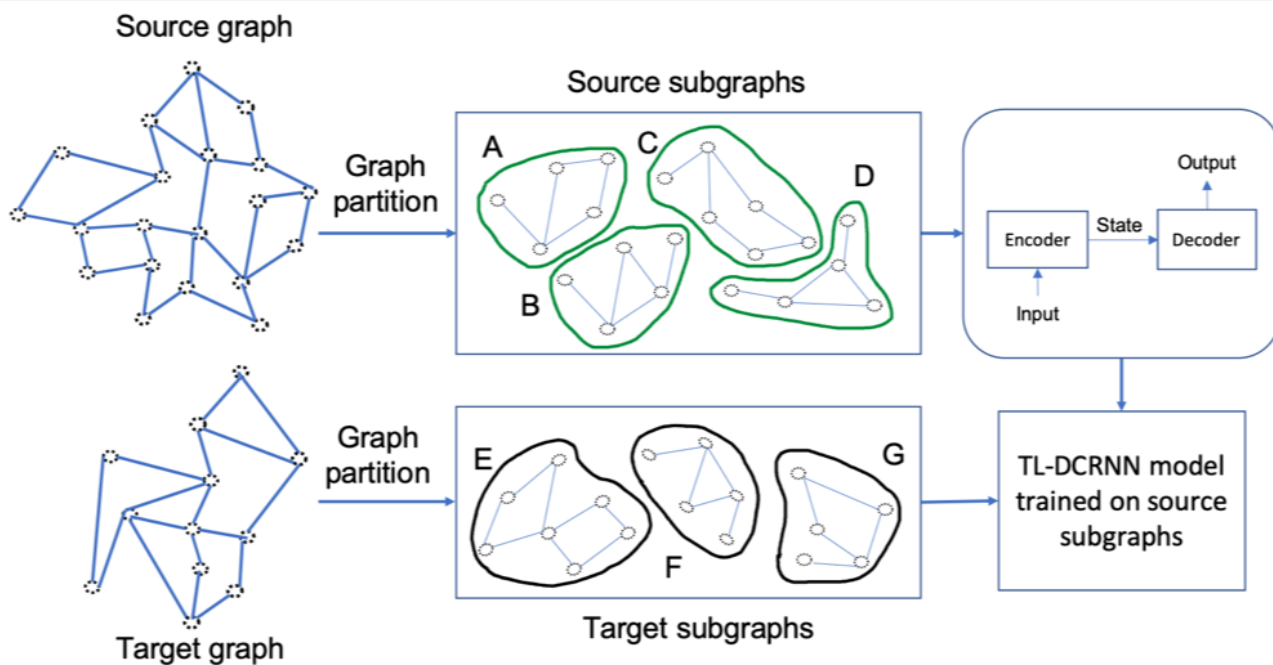
Bridge	Observed Count (Literature) <sup>1</sup>	Estimated Count (Mobility Baseline)	Relative Error	Dynamic Traffic Assignment			
				Bridge	Observed Count (Literature) <sup>1</sup>	Estimated Count (DTA)	Relative Error
Bay Bridge	247,500	238,648	-3.58%				
San Mateo	97,000	94,081	-3.01%	Bay Bridge	247,500	257,956	-3.58%
Dumbarton CA 84	81,800	97,218	18.85	San Mateo	97,000	108,643	-12%
Richmond	79,200	79,610	0.52	Dumbarton CA 84	81,800	105,921	-30%
Golden Gate	112,000	94,230	-15.87%	Richmond	79,200	89,500	-13%
				Golden Gate	112,000	104,144	-6%

Source: <sup>1</sup> Current and Projected Conditions Report San Francisco Bay Crossings Study Update Prepared for the Bay Area Toll Authority, 2010

# Validation: Detector Counts



# Transfer Learning with Graph Neural Networks (TL-DCRNN)



Method	MAE	RMSE	MAPE
Training and testing on PEMS-BAY			
ARIMA	3.38	6.50	8.30%
SVR	3.28	7.08	8.00%
FNN	2.46	4.98	5.89%
FC-LSTM	2.37	4.96	5.70%
STGCN	2.49	5.69	5.79%
DCRNN	2.07	4.74	4.90%
Graph Wevenet	1.95	4.52	4.63%
GMAN	1.86	4.32	4.31%
Training on LA and testing on PEMS-BAY			
TL-DCRNN	$2.13 \pm 1.09$	$5.23 \pm 2.29$	$5.55 \pm 4.34$

TL-DCRNN predicted speeds and flows better than many models.  
It was trained with Los Angeles data and tested on San Francisco network.

# Project Publications

- **Transfer Learning with Graph Neural Networks for Short-Term Highway Traffic Forecasting, Submitted to KDD 2020 : Tanwi Mallick, Prasanna Balaprakash, Eric Rask, Jane Macfarlane**
- **Graph-Partitioning-Based Diffusion Convolution Recurrent Neural Network for Large-Scale Traffic Forecasting, TRB 2020 : Tanwi Mallick, Prasanna Balaprakash, Eric Rask, Jane Macfarlane, Accepted for Transportation Research Record**
- **Mobiliti: Scalable Transportation Simulation Using High-Performance Parallel Computing, IEEE Intelligent Transportation Systems Conference ; Cy Chan, Bin Wang, John Bachan, and Jane Macfarlane**
- **Data-Driven Energy Use Estimation in Large Scale Transportation Networks, ACM Smart Cities Conference: Bin Wang, Cy Chan, Divya Somasi, Jane Macfarlane, Eric Rask**
- **Designing for Mode Shift Opportunity with Metropolitan Scale Simulation, ACM Smart Cities Conference: Kanaad Deodhar, Colin Laurence, Jane Macfarlane**
- **Assessing the Equity Implications of Localized Emissions Impacts From Transportation Using Dynamic Traffic Assignment A Case Study of the Los Angeles Region, Submitted to DTA 2020: Jessica Lazarus, Ioanna Kavvada, Ahmad Bin Thaneya, Bin Wang, and Jane Macfarlane**
- **Assessing the Equity Implications of Localized Congestion and Emissions Impacts of Four Traffic Assignment Scenarios in the Los Angeles Basin, TRB 2020: Jessica Lazarus, Ioanna Kavvada, Ahmad Bin Thaneya, Bin Wang, and Jane Macfarlane**
- **A traffic demand analysis method for Urban Air Mobility," by Bulusu, Vishwanath; Onak, Emin; Sengupta, Raja; Macfarlane, Jane , Submitted to Special Issue on Unmanned Aircraft System Traffic Management) IEEE Intelligent Transportation Systems Transactions.**